

DOCKET NO.: 272841US6PCT

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

IN RE APPLICATION OF: Mitsuhiro MIYAZAKI, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HERewith

INTERNATIONAL APPLICATION NO.: PCT/JP03/15927

INTERNATIONAL FILING DATE: December 12, 2003

FOR: INFORMATION PROCESSING DEVICE AND METHOD, RECORDING MEDIUM, AND PROGRAM

**REQUEST FOR PRIORITY UNDER 35 U.S.C. 119  
AND THE INTERNATIONAL CONVENTION**

Commissioner for Patents  
Alexandria, Virginia 22313

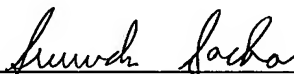
Sir:

In the matter of the above-identified application for patent, notice is hereby given that the applicant claims as priority:

<u>COUNTRY</u>	<u>APPLICATION NO</u>	<u>DAY/MONTH/YEAR</u>
Japan	2002-361540	12 December 2002
Japan	2003-285030	01 August 2003

Certified copies of the corresponding Convention application(s) were submitted to the International Bureau in PCT Application No. PCT/JP03/15927. Receipt of the certified copy(s) by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.

Respectfully submitted,  
OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.



Gregory J. Maier  
Attorney of Record  
Registration No. 25,599  
Surinder Sachar  
Registration No. 34,423

Customer Number

**22850**

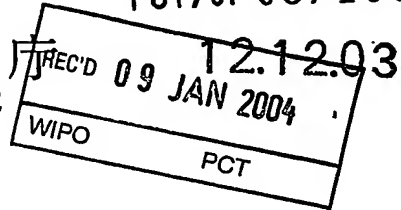
(703) 413-3000  
Fax No. (703) 413-2220  
(OSMMN 08/03)

Rec'd PCT/PTO 10 JUN 2005

10/538658

PCT/JP 03/15927

日本国特許  
JAPAN PATENT OFFICE



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application: 2002年12月12日

出願番号  
Application Number: 特願2002-361540  
[ST. 10/C]: [JP2002-361540]

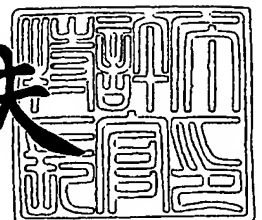
出願人  
Applicant(s): ソニー株式会社

PRIORITY  
DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2003年 8月25日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2003-3069508

【書類名】 特許願

【整理番号】 0290798804

【提出日】 平成14年12月12日

【あて先】 特許庁長官殿

【国際特許分類】 H04N 7/08

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 宮崎 充弘

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 山本 則行

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 斎藤 真里

【発明者】

【住所又は居所】 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社  
内

【氏名】 小池 宏幸

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代理人】

【識別番号】 100082131

【弁理士】

【氏名又は名称】 稲本 義雄

【電話番号】 03-3369-6479

【手数料の表示】

【予納台帳番号】 032089

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9708842

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報処理装置および方法、記録媒体、並びにプログラム

【特許請求の範囲】

【請求項 1】 配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、前記コンテンツを分類する分類手段と、

前記グループに属する前記コンテンツの利用履歴に基づいて、前記グループの利用頻度を算出する算出手段と、

前記算出手段により算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成手段と、

前記生成手段により生成された前記利用頻度に基づいて選択された前記グループに属するコンテンツを推薦する推薦手段と

を備えることを特徴とする情報処理装置。

【請求項 2】 前記分類手段は、前記コンテンツの放送時間帯と少なくとも 1 つ以上の他の属性に基づくグループに前記コンテンツを分類する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】 前記分類手段は、前記コンテンツを、少なくとも前記コンテンツの放送時間帯に基づくグループに分類するとともに、他の属性に基づくグループに分類し、

前記生成手段は、それぞれの前記グループの利用頻度を利用して、前記ユーザ嗜好情報を生成する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 4】 前記分類手段は、前記コンテンツの属性を形態素解析し、その結果得られた要素に基づくグループに、前記コンテンツを分類する

を特徴とする請求項 1 に記載の情報処理装置。

【請求項 5】 前記生成手段は、前記グループに属する前記コンテンツの利用状態が所定の条件を満たしている前記グループの利用頻度を利用して、前記ユーザ嗜好情報を生成する

ことを特徴とする請求項 1 に記載の情報処理装置。

【請求項 6】 配信されるコンテンツの属性を表す項目から生成された複数

のグループ化項目に基づくグループに、前記コンテンツを分類する分類ステップと、

前記グループに属する前記コンテンツの利用履歴に基づいて、前記グループの利用頻度を算出する算出ステップと、

前記算出ステップの処理で算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、

前記生成ステップの処理で生成された前記利用頻度に基づいて選択された前記グループに属するコンテンツを推薦する推薦ステップと  
を含むことを特徴とする情報処理方法。

【請求項 7】 配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、前記コンテンツを分類する分類ステップと、

前記グループに属する前記コンテンツの利用履歴に基づいて、前記グループの利用頻度を算出する算出ステップと、

前記算出ステップの処理で算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、

前記生成ステップの処理で生成された前記利用頻度に基づいて選択された前記グループに属するコンテンツを推薦する推薦ステップと  
を含むことを特徴とするコンピュータが読み取り可能なプログラムが記録されている記録媒体。

【請求項 8】 配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、前記コンテンツを分類する分類ステップと、

前記グループに属する前記コンテンツの利用履歴に基づいて、前記グループの利用頻度を算出する算出ステップと、

前記算出ステップの処理で算出された前記利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、

前記生成ステップの処理で生成された前記利用頻度に基づいて選択された前記グループに属するコンテンツを推薦する推薦ステップと

を含む処理をコンピュータに実行させることを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、情報処理装置および方法、記録媒体、並びにプログラムに関し、特に、コンテンツの推薦を効率よく、かつ効果的に行うことができる情報処理装置および方法、記録媒体、並びにプログラムに関する。

【0002】

【従来の技術】

配信されるコンテンツから、ユーザの嗜好の合うコンテンツを検出し、そのコンテンツの情報を提供するコンテンツ推薦システムが存在する（例えば、特許文献）。

【0003】

このシステムでは、例えば、ユーザがよく利用するコンテンツの属性（例えば、ジャンル）を検出し、その属性毎にコンテンツが推薦される。

【0004】

【特許文献】

特開 2000-287189 号公報

【0005】

【発明が解決しようとする課題】

しかしながら、コンテンツの属性は、コンテンツの編成等の事情に基づいてコンテンツの提供元により設定されているので、その属性毎の推薦では、推薦すべきコンテンツを効率的または効果的に検出することができない場合があった。

【0006】

本発明はこのような状況に鑑みてなされたものであり、コンテンツの推薦を行う側が、コンテンツの属性を利用してコンテンツをグループ化し、グループ毎にコンテンツの推薦を行うことができるようにするものである。

【0007】

【課題を解決するための手段】

本発明の情報処理装置は、配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、コンテンツを分類する分類手段と、グループに属するコンテンツの利用履歴に基づいて、グループの利用頻度を算出する算出手段と、算出手段により算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成手段と、生成手段により生成された利用頻度に基づいて選択されたグループに属するコンテンツを推薦する推薦手段とを備えることを特徴とする。

**【0008】**

分類手段は、コンテンツの放送時間帯と少なくとも1つ以上の他の属性に基づくグループにコンテンツを分類することを特徴とする。

**【0009】**

分類手段は、コンテンツを、少なくともコンテンツの放送時間帯に基づくグループに分類するとともに、他の属性に基づくグループに分類し、生成手段は、それぞれのグループの利用頻度を利用して、ユーザ嗜好情報を生成することができる。

**【0010】**

分類手段は、コンテンツの属性を形態素解析し、その結果得られた要素に基づくグループに、コンテンツを分類することができる。

**【0011】**

生成手段は、グループに属するコンテンツの利用状態が所定の条件を満たしているグループの利用頻度を利用して、ユーザ嗜好情報を生成することができる。

**【0012】**

本発明の情報処理方法は、配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、コンテンツを分類する分類ステップと、グループに属するコンテンツの利用履歴に基づいて、グループの利用頻度を算出する算出ステップと、算出ステップの処理で算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、生成ステップの処理で生成された利用頻度に基づいて選択されたグループに属するコンテンツを推薦する推薦ステップとを含むことを特徴とする。

## 【0013】

本発明のプログラムは、配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、コンテンツを分類する分類ステップと、グループに属するコンテンツの利用履歴に基づいて、グループの利用頻度を算出する算出ステップと、算出ステップの処理で算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報を生成する生成ステップと、生成ステップの処理で生成された利用頻度に基づいて選択されたグループに属するコンテンツを推薦する推薦ステップとを含むことを特徴とする。

## 【0014】

本発明の情報処理装置および方法、並びにプログラムにおいては、配信されるコンテンツの属性を表す項目から生成された複数のグループ化項目に基づくグループに、コンテンツが分類され、グループに属するコンテンツの利用履歴に基づいて、グループの利用頻度が算出され、算出された利用頻度に基づいて、ユーザの嗜好を表すユーザ嗜好情報が生成され、生成された利用頻度に基づいて選択されたグループに属するコンテンツが推薦される。

## 【0015】

## 【発明の実施の形態】

図1は、本発明を適用したコンテンツ推薦システムの構成例を示している。

## 【0016】

配信サーバ3は、ストリーミングデータデータベース1から、ストリーミングデータを取得し、インターネットその他のネットワークを含むネットワーク6を介して、クライアント機器5に配信する。配信サーバ3はまた、メタデータデータベース2から、コンテンツのメタデータを取得し、ネットワーク6を介して、コンテンツ推薦サーバ4に供給する。

## 【0017】

メタデータは、コンテンツ毎に、例えば、図2に示すような、「放送開始時刻」、「放送終了時刻」、「放送局」、「ジャンル」、「タイトル」、「出演者名」、「番組内容」、および「キーワード」等のコンテンツの属性を表す項目から構成されている。

**【0018】**

コンテンツ推薦サーバ4は、配信サーバ3から供給されたメタデータの所定の項目をグループ化項目とし、その項目の構成要素を所定のグループに分類し、グループ毎にIDを割り当てる。

**【0019】**

例えば図3に示すように、メタデータの項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」が1つのグループ化項目とされ、その構成要素である「8ch、10ch、・・・」と「0:00～24:00」が、各放送局と1時間の放送時間の組にグループ化され、そのグループ毎にIDが割り当てられる。

**【0020】**

また、図4に示すように、メタデータの項目「ジャンル」および項目「出演者」が1つのグループ化項目とされ、その構成要素である「バラエティ、ニュース、映画、・・・」と「タレントA、タレントB、・・・」が、各ジャンルと各出演者の組にグループ化され、そのグループ毎にIDが割り当てられる。

**【0021】**

なお、図3、4の例では、グループ化項目が、メタデータの複数の項目から構成されているが、1つの項目とすることもできる。

**【0022】**

コンテンツ推薦サーバ4はまた、図5に示すように、メタデータのそれぞれに、そのメタデータの項目の内容に対応するグループのグループIDを設定する。なおコンテンツの属性によっては、図5に示すように、複数のグループIDが設定される場合もある。

**【0023】**

コンテンツ推薦サーバ4は、このようにグループIDが設定されたメタデータを、適宜、クライアント機器5に送信する。

**【0024】**

コンテンツ推薦サーバ4はまた、クライアント機器5から、コンテンツのグループIDを含む利用履歴を適宜取得し、その利用履歴に基づいて、グループの利

用頻度を算出する。そしてコンテンツ推薦サーバ4は、算出したその利用頻度をユーザの嗜好を表すものとして利用し、グループ毎にコンテンツの推薦を行う。詳細は後述するが、高い利用頻度のグループに属するコンテンツに関する情報が、コンテンツ推薦情報としてクライアント機器5に送信される。

#### 【0025】

クライアント機器5は、配信サーバ3から配信されてきたコンテンツを利用するが、その利用履歴として、利用したコンテンツの、例えば図6に示すようなメタデータ（グループIDが設定されたもの）を、適宜、コンテンツ推薦サーバ4に供給する。

#### 【0026】

クライアント機器5は、コンテンツ推薦サーバ4から供給されたコンテンツ推薦情報を表示してユーザに提示する。ユーザは、それを参照することにより、自分の嗜好にあったコンテンツを選択することができる。

#### 【0027】

なお、ここでの配信サーバ3乃至クライアント機器5の通信は、ネットワーク6を介して行われているが、それぞれ直接通信する構成にすることもできる。

#### 【0028】

図7は、コンテンツ推薦サーバ4の構成例を示している。CPU（Central Processing Unit）11は、ROM（Read Only Memory）12に記憶されている、例えば、コンテンツ推薦用のプログラム等に従って所定の処理を実行する。RAM（Random Access Memory）13には、CPU11がその処理を実行する上において必要なデータなどが適宜記憶される。

#### 【0029】

CPU11にはバス14を介して入出力インタフェース15が接続されている。入出力インタフェース15には、キーボード、マウスなどよりなる入力部16、LCD（Liquid Crystal Display）などよりなる出力部17、メタデータ等を記憶する記憶部18、およびネットワーク6を介して配信サーバ3またはクライアント機器5との通信を行う通信部19が接続されている。

#### 【0030】

入出力インタフェース 15 には、ドライブ 20 が適直接続され、CPU 11 は、そこに装着される磁気ディスク 31、光ディスク 32、光磁気ディスク 33、または半導体メモリ 34 との間でデータの授受を行う。

#### 【0031】

図 8 は、クライアント機器 5 の構成例を示している。この構成は、コンテンツ推薦サーバ 4 の構成と基本的に同様であるので、その説明は省略する。

#### 【0032】

次に、ユーザ嗜好情報を生成する場合のコンテンツ推薦サーバ 4 の動作を、図 9 のフローチャートを参照して説明する。

#### 【0033】

ステップ S1 において、コンテンツ推薦サーバ 4 の CPU 11 は、ユーザ嗜好情報を生成するタイミングであるか否かを判定し、そのタイミングであると判定した場合、ステップ S2 に進む。例えば、クライアント機器 5 からコンテンツ推薦情報（後述）の提供が要求されたとき、または予め決められた時期（例えば、毎週所定の時刻）が来たとき、ステップ S2 に進む。

#### 【0034】

ステップ S2 において、CPU 11 は、通信部 19 を介して、クライアント機器 5 から、所定の利用履歴を取得する。この例の場合、1 週間前からの間に利用されたコンテンツのメタデータ（グループ ID が設定されている）が取得される。CPU 11 は、グループ化項目毎に、グループ毎の利用頻度を算出する。

#### 【0035】

この例の場合、図 3 に示したように項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目がグループ化されているので、図 10 に示すように、そのグループ毎の利用頻度（回数）が算出される。なお、図 10 には、利用回数の高い上位 3 個のグループが示されている。

#### 【0036】

また図 4 に示したように項目「ジャンル」および項目「出演者」からなるグループ化項目もグループ化されているので、図 11A に示すように、そのグループ毎の利用回数が算出される。図 11A にも、利用回数が高い上位 3 個のグループ

が示されている。

#### 【0037】

なお、図11Aの利用回数は、コンテンツの配信数に応じて多くなる可能性があるので、そのままではユーザの嗜好を正確に対応しない。そのため、ステップS2で取得された利用履歴に対応する期間中に配信されたコンテンツ数で利用回数が正規化される。

#### 【0038】

例えば、タレントDが出演するバラエティ番組が、その期間（この例の場合、1週間）の間に10本配信され、タレントDが出演するニュース番組が、その期間の間に100本配信され、タレントCが出演するバラエティ番組が、その期間の間に80本配信された場合、図11Aの利用回数は、図11Bに示すように、正規化される。このように利用回数を正規化することでユーザの嗜好を適切に対応した利用頻度を得ることができる。

#### 【0039】

図9に戻り、ステップS3において、コンテンツ編集サーバ4のCPU11は、グループ化項目毎に、所定の閾値以上の利用頻度が得られたグループを検出する。

#### 【0040】

例えば、項目「放送局」、項目「放送開始時刻」、および項目「放送開始時刻」からなるグループ化項目の閾値が値7である場合、図10の例では、グループ「8ch, 20:00~21:00」、およびグループ「10ch, 19:00~20:00」の2つのグループが検出される。

#### 【0041】

また項目「ジャンル」および項目「出演者」からなるグループ化項目の閾値が値0.06である場合、図11Bの例では、グループ「バラエティ, タレントD」、グループ「ニュース, タレントD」、およびグループ「バラエティ, タレントC」の3つのグループが検出される。

#### 【0042】

次に、ステップS4において、CPU11は、ステップS3で検出したグループ

に属するコンテンツの利用状態が所定の条件を満たしているか否かを判定し、その条件を満たしている場合、そのグループを、ユーザに嗜好に対応するものとして、後述するユーザ嗜好情報の生成に利用する。一方、その条件を満たしていない場合、そのグループは、ユーザ嗜好情報の生成に利用されない。

#### 【0043】

具体的には、そのグループに属するコンテンツの配信リストに基づいて、いまから遡って所定の回数連続して利用されなかったか否かが判定される。

#### 【0044】

例えば、その所定の回数が3回である場合、図12に示すように、グループ「8ch, 20:00~21:00」に属する番組の、いまから遡って最も最近の番組Aは視聴されなかったが、その前に配信された番組Bは視聴されているとき、グループ「8ch, 20:00~21:00」は、ユーザの嗜好に対応するものとされる。

#### 【0045】

一方、図13に示すように、グループ「バラエティ, タレントD」に属するコンテンツが3回連続して視聴されなかったとき、グループ「バラエティ, タレントD」は、ユーザの嗜好に対応しないものとされる。

#### 【0046】

なお、図14に示すように、所定の回数（例えば、3回）連続して利用されたか否かを条件としてユーザの嗜好に対応するグループを選ぶこともできる。

#### 【0047】

ステップS6において、CPU11は、ステップS5で、ユーザの嗜好に対応するものとされたグループのグループIDをユーザ嗜好情報として、記憶部18に記憶する。

#### 【0048】

この例の場合、項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目におけるグループ「8ch, 20:00~21:00」とグループ「10ch, 19:00~20:00」のグループID、および項目「ジャンル」および項目「出演者」からなるグループ化項目におけるグループ

「ニュース，タレント D」とグループ「バラエティ，タレント C」のグループ ID が、ユーザ嗜好情報として記憶部 1 8 に記憶される。

【 0 0 4 9 】

次に、コンテンツ推薦情報を生成する場合のコンテンツ推薦サーバ 4 の動作を、図 1 5 のフローチャートを参照して説明する。

【 0 0 5 0 】

ステップ S 2 1 において、コンテンツ推薦サーバ 4 の CPU 1 1 は、クライアント機器 5 から、コンテンツ推薦情報の提供が要求されるまで待機し、その要求があったとき、ステップ S 2 2 に進み、記憶部 1 8 から上述したようにして生成したユーザ嗜好情報を取得する。

【 0 0 5 1 】

ステップ S 2 3 において、CPU 1 1 は、いまから配信されるコンテンツのメタデータ（グループ ID が設定されているもの）から、ユーザ嗜好情報としてのグループ ID と同じグループ ID が設定されているメタデータを抽出する。CPU 1 1 は、抽出したメタデータからコンテンツ推薦情報を生成する。

【 0 0 5 2 】

なお、グループ化項目が複数ある場合には（例えば、図 3，4）には、各グループ化項目から抽出されたグループのいずれにも属するコンテンツのメタデータを抽出することもできる。

【 0 0 5 3 】

ステップ S 2 4 において、CPU 1 1 は、通信部 1 9 を介して、ステップ S 2 3 で生成したコンテンツ推薦情報を、クライアント機器 5 に送信する。クライアント機器 5 は、コンテンツ推薦サーバ 4 から送信されてきたコンテンツ推薦情報を出力部 5 7 に表示する。

【 0 0 5 4 】

図 1 6 および図 1 7 は、コンテンツ推薦情報の表示例を示している。

【 0 0 5 5 】

図 1 6 の例では、項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」からなるグループ化項目のグループ「8 ch，2 0：0 0～2 1：0 0

」に属する番組AAAの情報（タイトル等）と、グループ「10ch, 19:00～20:00」に属する番組BBAの情報（タイトル等）が表示されている。

#### 【0056】

図17の例では、項目「ジャンル」および項目「出演者」からなるグループ化項目のグループ「ニュース, タレントD」およびグループ「バラエティ, タレントC」に属する番組に関する情報（タイトル等）が表示されている。なお、図17の例の場合、それぞれのグループに属する番組の情報が表示されるウィンドウが一部重なって表示されている。

#### 【0057】

ユーザは、このように表示されたコンテンツ推薦情報を参照して、視聴する番組を選択することができる。

#### 【0058】

以上のように、ユーザの嗜好を把握するための利用頻度を、グループIDを利用してグループ毎に算出するようにしたので、メタデータの項目毎に利用頻度を算出する場合に比べ、その計算量を少なくすることができる。

#### 【0059】

また、コンテンツ推薦情報がグループ毎にまとめて表示されるようにしたので、表示スペースが小さいクライアント機器5においても適切にコンテンツ推薦情報を表示することができる。

#### 【0060】

また、以上においては、メタデータの項目「放送局」、項目「放送開始時刻」、および項目「放送終了時刻」、並びに項目「ジャンル」および項目「出演者」を利用してグループ化を行ったが、項目「タイトル」や項目「内容」など他の項目を利用してグループ化を行うこともできる。その結果、例えば、再放送やスペシャル版の番組をオリジナルの番組と同じグループに属するコンテンツとして扱うことができるので、オリジナルであろうと、再放送されたものでであろうと、その番組が視聴されれば、その利用履歴をユーザ嗜好情報生成に反映することができる。

#### 【0061】

ここで、図18のフローチャートを参照して、項目「タイトル」を利用してグループ化を行う処理について説明する。

【0062】

ステップS61において、コンテンツ推薦サーバ4は、メタデータから、タイトルを抽出する。

【0063】

ステップS62において、コンテンツ推薦サーバ4は、タイトルを形態素解析し、単語に分解する。具体的には、メタデータに含まれている映画の題名が「東海道三谷怪談」であった場合、これがタイトルとして形態素解析されるので、「東海道」、「三谷」、「怪談」の3つの単語が得られる。

【0064】

ステップS63において、コンテンツ推薦サーバ4は、解析された単語、もしくは、複数の単語から構成される単語群のうちのいずれかを抽出して、記憶部18から、抽出された単語、または単語群に対応するグループIDを抽出する。

【0065】

ここで、複数の単語から構成される単語群とは、形態素解析により得られた単語の組み合わせにより生成される単語群であり、例えば、形態素解析により得られた単語が「東海道」、「三谷」、「怪談」である場合、単語群は、「東海道・三谷」、「東海道・怪談」、「三谷・怪談」となる。

【0066】

ステップS64において、コンテンツ推薦サーバ4は、グループIDが抽出されたか否かを判断する。

【0067】

ステップS64において、対応するグループIDが抽出されなかったと判断された場合、抽出された単語、もしくは、複数の単語から構成される単語群には、まだグループIDが付けられていないので、ステップS65において、抽出された単語、もしくは、複数の単語から構成される単語群に新たなグループIDを対応付ける。また、コンテンツ推薦サーバ4は、単語、もしくは、複数の単語から構成される単語群と、それに対応するグループIDを記憶する。

## 【0068】

ステップS64において、対応するグループIDが抽出されたと判断された場合、または、ステップS65の処理の終了後、ステップS66において、コンテンツ推薦サーバ4は、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループIDを抽出したか否かを判断する。

## 【0069】

ステップS66において、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループIDを抽出していないと判断された場合、処理は、ステップS63に戻り、それ以降の処理が繰り返される。

## 【0070】

ステップS66において、タイトルを構成する全ての単語、もしくは、複数の単語から構成される単語群についてグループIDが抽出されたと判断された場合、ステップS67において、コンテンツ推薦サーバ4は、メタデータに、抽出されたグループIDを対応付けて、処理が終了される。

## 【0071】

なお、類似したタイトルの番組を、同一のグループとするようにしても良い。例えば、タイトル「2年A組銀八先生」の連続ドラマと、タイトル「2年A組銀八先生スペシャル」の特別番組とを、同一のグループとしてグループ化することができるよう、タイトルを構成する単語の形態素解析結果を基に、例えば、2週間、1ヶ月、半年などの所定の期間の番組タイトルで、単語の一致率を総当りで算出し、単語の一致率が所定の値以上である場合、同一グループとするようにしてもよい。

## 【0072】

次に、図19のフローチャートを参照して、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理2について説明する。

## 【0073】

ステップS401およびステップS402において、図18を用いて説明した、ステップS61およびステップS62と同様の処理が実行される。すなわち、コンテンツ推薦サーバ4は、メタデータから、タイトルを抽出して形態素解析し

、単語に分解する。

【0074】

ステップS403において、コンテンツ推薦サーバ4は、解析された単語を基に、タイトル間の単語の一致度、すなわち、単語が一致している割合を示す一致率を算出する。

【0075】

具体的には、タイトル「2年A組銀八先生」と、タイトル「2年A組銀八先生スペシャル」とが、それぞれ、「2」「年」「A」「組」「銀八」「先生」と、「2」「年」「A」「組」「銀八」「先生」「スペシャル」とに形態素分析された場合、この2つの番組のタイトルを構成する単語の一致率は、6/7で85.7%となる。

【0076】

ステップS404において、コンテンツ推薦サーバ4は、単語が、例えば、70%などの所定の値以上一致しているか否かを判断する。この、一致率の閾値は、70%以外のいかなる数値であっても良いことは言うまでもない。

【0077】

ステップS404において、単語が、70%などの所定の値以上一致していると判断された場合、ステップS405において、コンテンツ推薦サーバ4は、それらの番組に、同一のグループIDを対応付ける。また、コンテンツ推薦サーバ4は、一致した単語、または、単語群と、それに対応するグループIDを記憶する。

【0078】

ステップS404において、70%などの所定の値以下の一致率であると判断された場合、または、ステップS405の処理の終了後、ステップS406において、コンテンツ推薦サーバ4は、タイトルの総当りが終了したか否かを判断する。

【0079】

ステップS406において、タイトルの総当りが終了していないと判断された場合、処理は、ステップS403に戻り、それ以降の処理が繰り返される。

**【0080】**

ステップS406において、タイトルの総当りが終了したと判断された場合、処理が終了される。

**【0081】**

このような処理により、タイトルを構成する単語の一致率を基にしたグループIDが対応付けられるので、例えば、連続ドラマとスペシャル番組などの類似したタイトルの番組を、同一のグループとして処理させるようにすることができる。

**【0082】**

また、タイトルを構成する単語の一致率を基にグループを決定するようにすることにより、例えば、メタデータにおいて、数字の半角と全角、または、英字の半角と全角、もしくは、大文字と小文字などの表記ゆれがあった場合にも、同一タイトルの番組を、同一のグループとして検出することが可能となる。

**【0083】**

また、単語の一致率に加えて、例えば、放送局や番組ジャンル、あるいは、放送開始時刻などを、グループ化の条件に加えるようにしても良い。例えば、ニュース番組などにおいては、タイトルが、「ニュース」を含む少ない単語によって構成されているので、図19を用いて説明した処理では、異なる放送局の異なる形態のニュース番組であっても、同一のグループとして検出されてしまう恐れがあるので、単語の一致率に加えて、放送局も一致した場合、同一グループとするようにしても良い。

**【0084】**

次に、図20のフローチャートを参照して、放送局の一致を条件に加えて、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理3について説明する。

**【0085】**

ステップS421乃至ステップS424において、図19を用いて説明した、ステップS401乃至ステップS404と同様の処理が実行される。すなわち、コンテンツ推薦サーバ4は、メタデータから、タイトルを抽出して形態素解析し

、単語に分解する。そして、コンテンツ推薦サーバ4は、解析された単語を基に、タイトル間の単語の一致度を算出し、単語が、例えば、70%などの所定の値以上一致しているか否かを判断する。

**【0086】**

ステップS424において、単語が、70%などの所定の値以上一致していると判断された場合、ステップS425において、コンテンツ推薦サーバ4は、それらの番組の放送局が一致しているか否かを判断する。

**【0087】**

ステップS425において、これらの番組の放送局が一致していると判断された場合、ステップS426において、コンテンツ推薦サーバ4は、それらの番組に、同一のグループIDを対応付ける。また、コンテンツ推薦サーバ4は、一致した単語、または、単語群と、それに対応する放送局およびグループIDを記憶する。

**【0088】**

ステップS424において、70%などの所定の値以下の一致率であると判断された場合、ステップS425において、これらの番組の放送局が一致していないと判断された場合、または、ステップS426の処理の終了後、ステップS427において、コンテンツ推薦サーバ4は、タイトルの総当りが終了したか否かを判断する。

**【0089】**

ステップS427において、タイトルの総当りが終了していないと判断された場合、処理は、ステップS423に戻り、それ以降の処理が繰り返される。

**【0090】**

ステップS427において、タイトルの総当りが終了したと判断された場合、処理が終了される。

**【0091】**

このような処理により、放送局の一致とタイトルを構成する単語の一致率を基にしたグループIDが対応付けられるので、例えば、類似したタイトルの番組を、同一のグループとする場合に、他局のニュース番組を同一のグループとするよ

うなことを防ぐことができる。

#### 【0092】

なお、図20においては、タイトルを構成する単語の一致率以外に、同一の放送局であるか否かを条件として、グループ化を行うものとして説明したが、放送局以外の、例えば、放送時間帯やジャンルなどを、タイトルを構成する単語の一致率以外の条件として、グループ化を実行するようにしても良いことは言うまでもない。

#### 【0093】

更に、例えば、連続ドラマや帯番組の放送開始時刻が、スポーツ中継や特別番組などのためにずれた場合においても、同一グループとして検出可能なように、タイトルを構成する単語の一致率以外の条件を、放送時刻が、例えば、1時間などの所定の時間範囲内のずれで一致しているか否かとして、グループ化を実行するようにしても良い。

#### 【0094】

図21のフローチャートを参照して、放送時刻が、所定の時間範囲内のずれで一致しているか否かを条件に加えて、タイトルを構成する単語の一致率によりグループ化を実行するタイトルグループ化処理4について説明する。

#### 【0095】

ステップS441乃至ステップS444において、図19を用いて説明した、ステップS401乃至ステップS404と同様の処理が実行される。すなわち、コンテンツ推薦サーバ4は、メタデータから、タイトルを抽出して形態素解析し、単語に分解する。そして、コンテンツ推薦サーバ4は、解析された単語を基に、タイトル間の単語の一致度を算出し、単語が、例えば、70%などの所定の値以上一致しているか否かを判断する。

#### 【0096】

ステップS444において、単語が、70%などの所定の値以上一致していると判断された場合、ステップS445において、コンテンツ推薦サーバ4は、これらの番組の放送開始時刻が、例えば、1時間などの所定の範囲内のずれで一致しているか否かを判断する。

## 【0097】

ステップS445において、それらの番組の放送開始時刻が所定の範囲内のずれで一致していると判断された場合、ステップS446において、コンテンツ推薦サーバ4は、それらの番組に、同一のグループIDを対応付ける。また、コンテンツ推薦サーバ4は、一致した単語、または、単語群と、それに対応する放送開始時刻の範囲、およびグループIDを記憶する。

## 【0098】

ステップS444において、70%などの所定の値以下の一致率であると判断された場合、ステップS445において、それらの番組の放送開始時刻が所定の範囲以上にずれていると判断された場合、または、ステップS446の処理の終了後、ステップS447において、コンテンツ推薦サーバ4は、タイトルの総当りが終了したか否かを判断する。

## 【0099】

ステップS447において、タイトルの総当りが終了していないと判断された場合、処理は、ステップS443に戻り、それ以降の処理が繰り返される。

## 【0100】

ステップS447において、タイトルの総当りが終了したと判断された場合、処理が終了される。

## 【0101】

このような処理により、放送開始時刻の所定の範囲内のずれを含む一致と、タイトルを構成する単語の一致率を基にしたグループIDが対応付けられるので、例えば、類似したタイトルの番組を同一のグループとする場合に、特別番組などによる放送時刻の変更のために、同一グループとして検出されるべき番組が、同一グループとして検出されないようなことを防ぐことができる。

## 【0102】

なお、以上においては、コンテンツ推薦サーバ4が、ユーザ嗜好情報生成処理(図9)およびコンテンツ推薦情報処理(図15)を行う場合を例として説明したが、クライアント機器5が、コンテンツ推薦サーバ4から供給されるグループIDが設定されたメタデータ(グループ化情報)を利用して、自分自身がグルー

ブ毎の利用頻度を算出してユーザ嗜好情報を生成し、それに基づいてコンテンツ推薦情報を生成することもできる。

#### 【0103】

上述した一連の処理は、ソフトウェアにより実行することもできる。そのソフトウェアは、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、記録媒体からインストールされる。

#### 【0104】

この記録媒体は、図7、8に示すように、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク31または71（フレキシブルディスクを含む）、光ディスク32または72（CD-ROM（Compact Disk-Read Only Memory）、DVD（Digital Versatile Disk）を含む）、光磁気ディスク33または73（MD（Mini-Disk）（商標）を含む）、もしくは半導体メモリ34または74などよりなるパッケージメディアなどにより構成される。

#### 【0105】

また、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

#### 【0106】

なお、本明細書において、システムとは、複数の装置により構成される装置全体を表すものである。

#### 【0107】

##### 【発明の効果】

本発明によれば、コンテンツの属性を表す項目から生成されたグループ化項目におけるグループ毎の利用頻度から、コンテンツ推薦を行うことができる。

##### 【図面の簡単な説明】

##### 【図1】

本発明を適用したコンテンツ推薦システムの構成例を示している。

【図 2】

メタデータの例を示す図である。

【図 3】

コンテンツのグループ化を説明する図である。

【図 4】

コンテンツのグループ化を説明する他の図である。

【図 5】

グループIDが付されたメタデータの例を示す図である。

【図 6】

利用履歴の例を示す図である。

【図 7】

図 1 のコンテンツ推薦サーバの構成例を示すブロック図である。

【図 8】

図 1 のクライアント機器の構成例を示すブロック図である。

【図 9】

ユーザ嗜好情報生成処理を説明するフローチャートである。

【図 10】

利用頻度の算出方法を説明する図である。

【図 11】

利用頻度の算出方法を説明する他の図である。

【図 12】

利用状態の確認処理を説明する図である。

【図 13】

利用状態の確認処理を説明する他の図である。

【図 14】

利用状態の確認処理を説明する他の図である。

【図 15】

コンテンツ推薦情報生成処理を説明するフローチャートである。

**【図 16】**

コンテンツ推薦情報の表示例を示す図である。

**【図 17】**

他のコンテンツ推薦情報の表示例を示す図である。

**【図 18】**

タイトルグループ化処理 1 について説明するフローチャートである。

**【図 19】**

タイトルグループ化処理 2 について説明するフローチャートである。

**【図 20】**

タイトルグループ化処理 3 について説明するフローチャートである。

**【図 21】**

タイトルグループ化処理 4 について説明するフローチャートである。

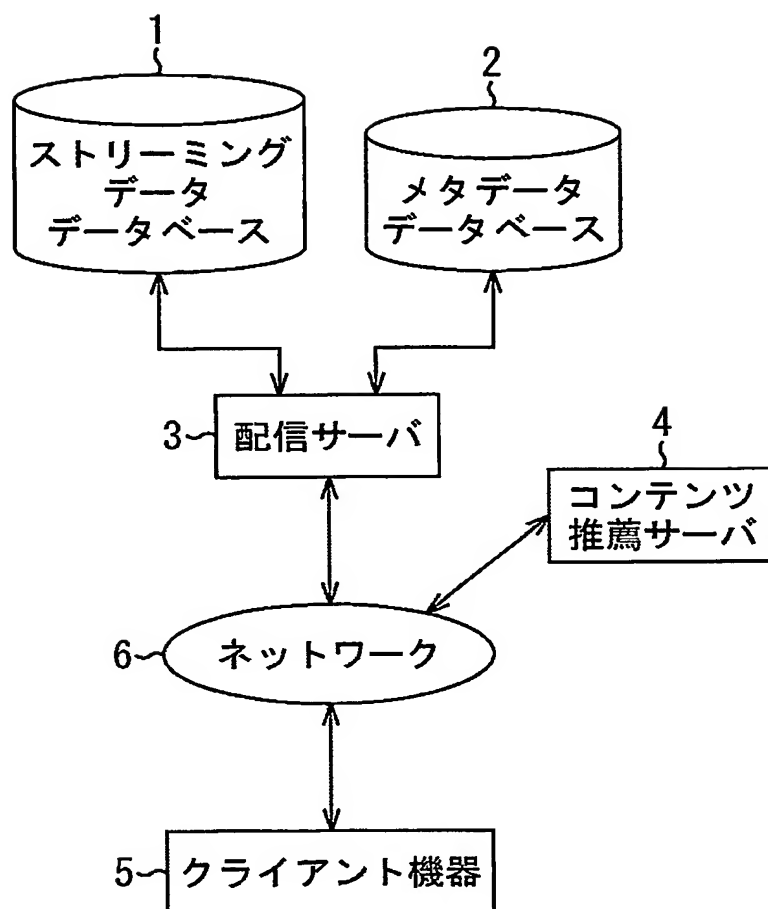
**【符号の説明】**

4 コンテンツ推薦サーバ, 5 クライアント機器, 11 CPU, 12 ROM, 13 RAM, 14 バス, 15 入出力インタフェース, 16 入力部, 17 出力部, 18 記憶部, 19 通信部

【書類名】 図面

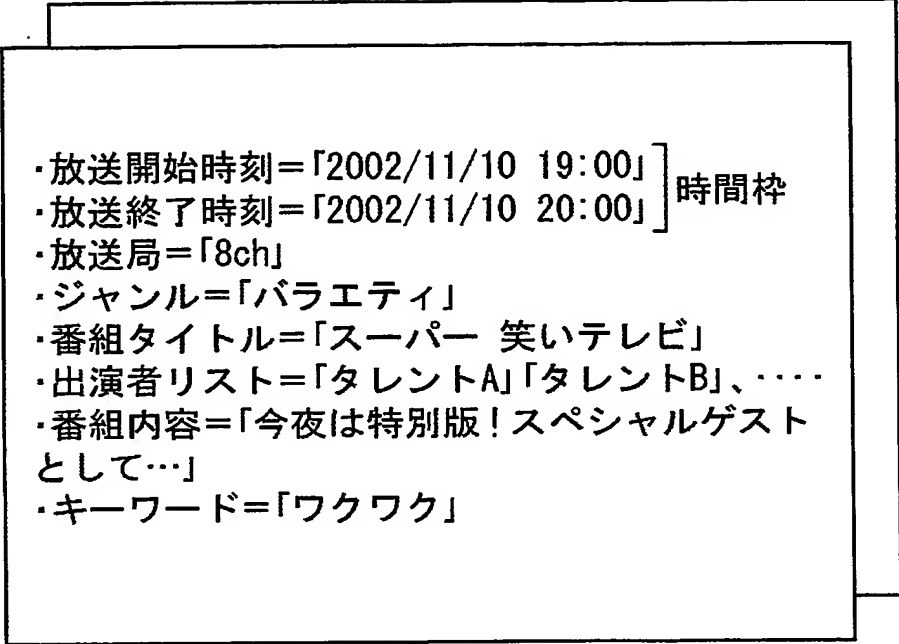
【図 1】

図1



【図2】

図2

- 
- ・放送開始時刻=「2002/11/10 19:00」
  - ・放送終了時刻=「2002/11/10 20:00」
  - ・放送局=「8ch」
  - ・ジャンル=「バラエティ」
  - ・番組タイトル=「スーパー 笑いテレビ」
  - ・出演者リスト=「タレントA」「タレントB」、……
  - ・番組内容=「今夜は特別版!スペシャルゲストとして…」
  - ・キーワード=「ワクワク」

【図3】

図3

「放送局」,「開始時刻」,「終了時刻」	グループID
8ch, 00:00~06:00	××××
8ch, 06:00~07:00	××××
8ch, 07:00~08:00	××××
⋮	
10ch, 00:00~06:00	××××
10ch, 06:00~07:00	××××
10ch, 07:00~08:00	××××
⋮	

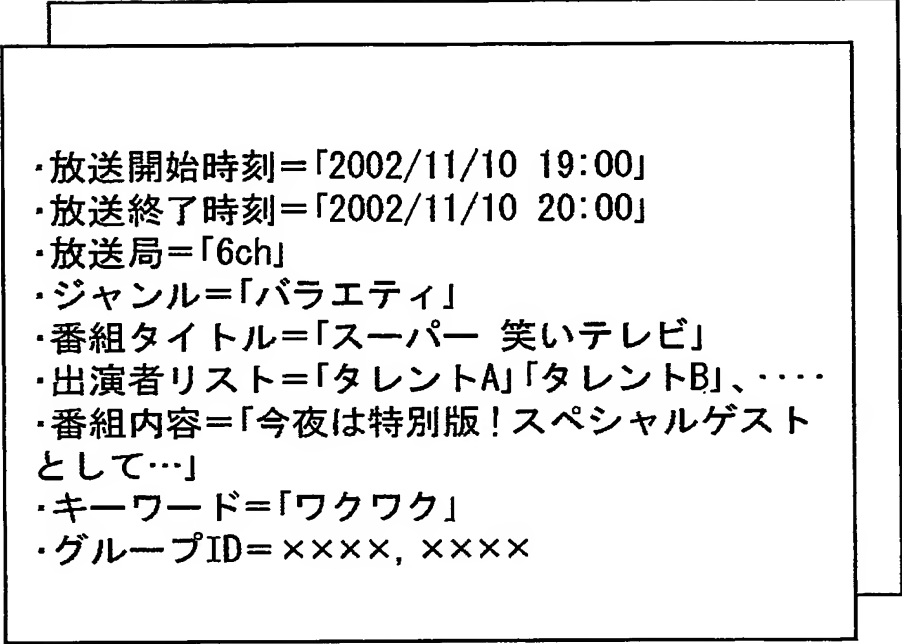
【図 4】

図4

「ジャンル」, 「出演者」	グループID
バラエティ, タレントA	××××
ニュース, タレントA	××××
映画, タレントA	××××
⋮	
バラエティ, タレントB	××××
ニュース, タレントB	××××
映画, タレントB	××××
⋮	

【図5】

図5

- 
- ・放送開始時刻=「2002/11/10 19:00」
  - ・放送終了時刻=「2002/11/10 20:00」
  - ・放送局=「6ch」
  - ・ジャンル=「バラエティ」
  - ・番組タイトル=「スーパー 笑いテレビ」
  - ・出演者リスト=「タレントA」「タレントB」、……
  - ・番組内容=「今夜は特別版!スペシャルゲストとして…」
  - ・キーワード=「ワクワク」
  - ・グループID=「xxxx, xxxx」

【図 6】

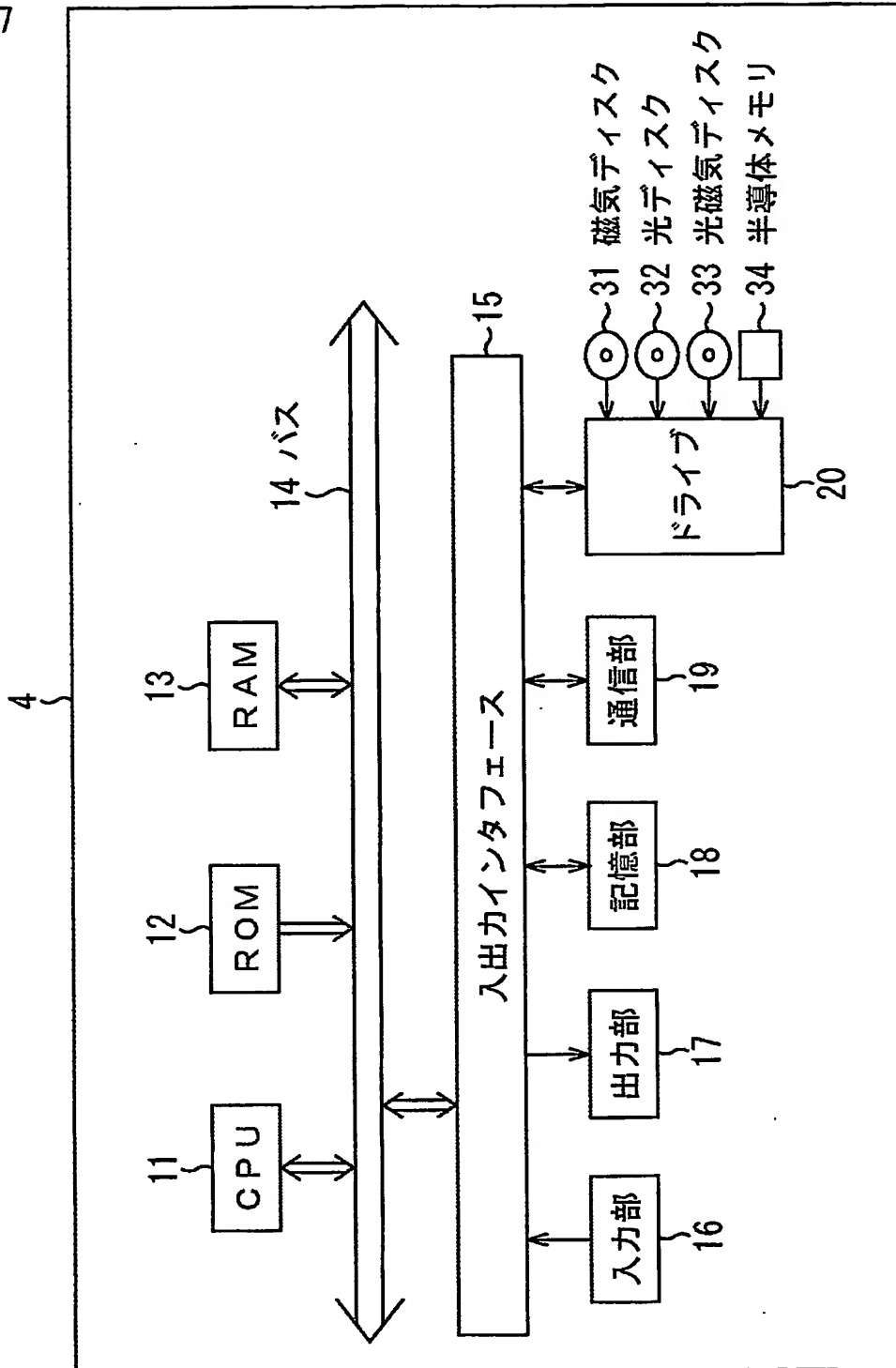
図6

- ・放送開始時刻=「2002/10/18 7:00」
- ・放送終了時刻=「2002/10/18 8:00」
- ・放送局=「8ch」
- ・ジャンル=「ニュース」
- ・番組タイトル=「モーニングニュース」
- ・出演者リスト=「タレントC」「タレントD」、……
- ・番組内容=「ニュース」
- ・キーワード=「××××」

グループID=××××, ××××

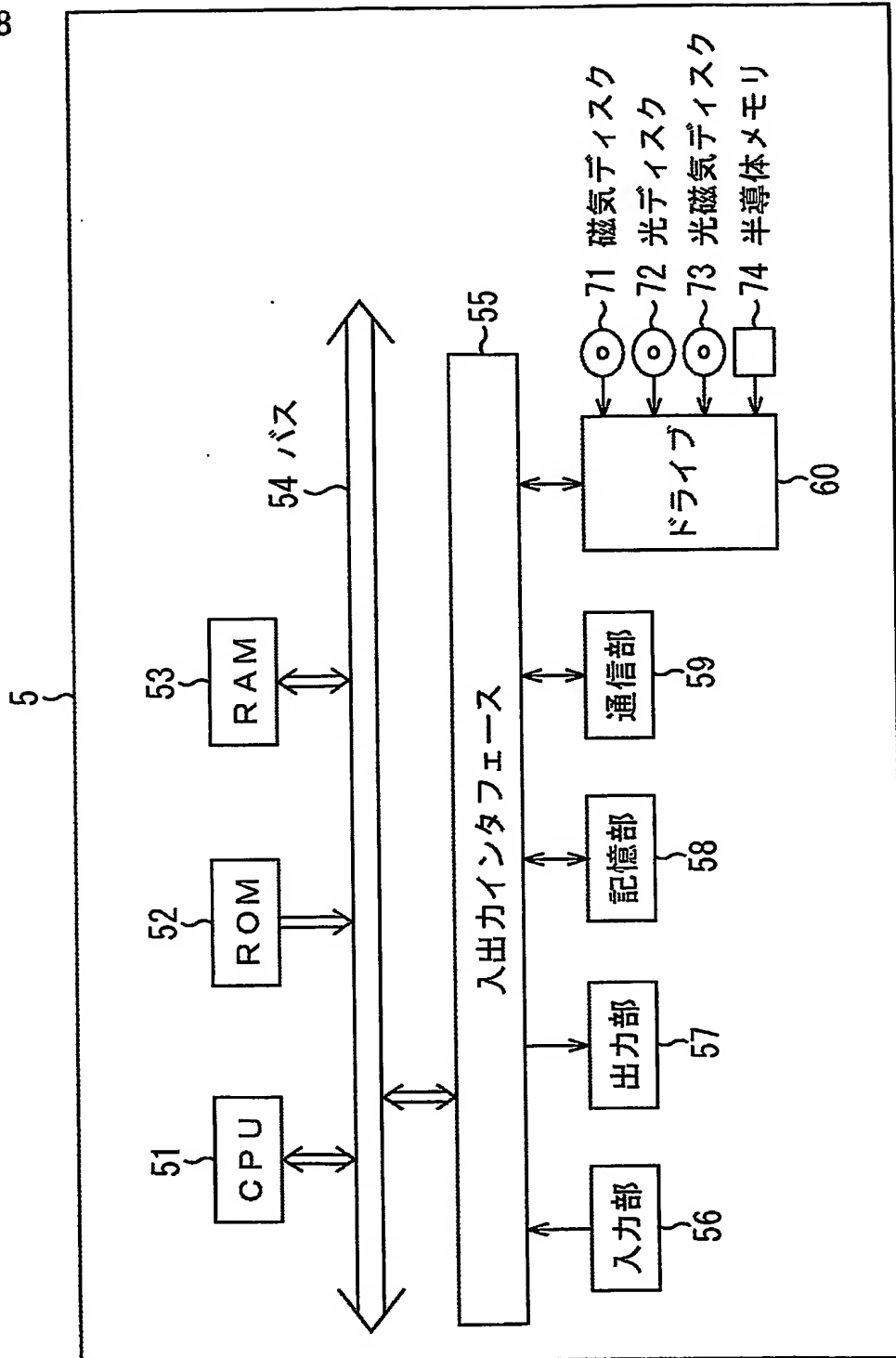
【図 7】

図 7



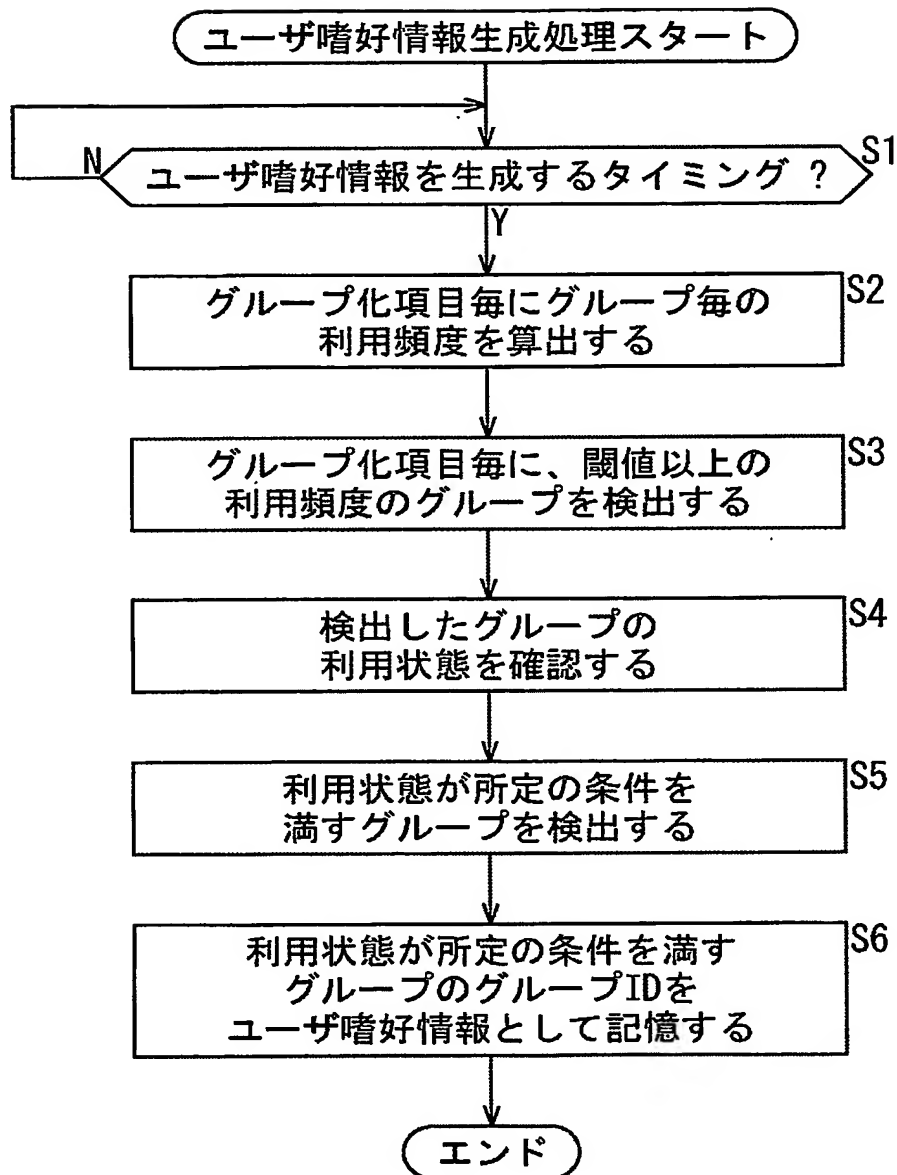
【図 8】

図8



【図 9】

図9



【図 1 0】

図10

「放送局」,「開始時刻」,「終了時刻」	回数
8ch, 20:00～21:00	7
10ch, 19:00～20:00	7
8ch, 22:00～23:00	6
⋮	

【図 11】

図11

「ジャンル」, 「出演者」	正規化回数
バラエティ, タレントD	0.25
ニュース, タレントD	0.08
バラエティ, タレントC	0.0625
...	

B

「ジャンル」, 「出演者」	回数
バラエティ, タレントD	10
ニュース, タレントD	8
バラエティ, タレントC	5
...	

A

【図 12】

図12

配信リスト 「8ch, 20:00~21:00」	利用履歴 「8ch, 20:00~21:00」
<div data-bbox="500 1402 548 1556">10/25(土)</div> <div data-bbox="500 1192 548 1297">番組 A</div> <div data-bbox="574 1402 623 1556">10/24(金)</div> <div data-bbox="574 1192 623 1297">番組 B</div>	<div data-bbox="500 594 548 877">視聴されていない</div> <div data-bbox="574 699 623 877">視聴された</div>

【図 13】

図13

配信リスト「バラエティ、タレントD」	利用履歴「バラエティ、タレントD」
<div data-bbox="500 1402 542 1556">10/25(土)</div> <div data-bbox="500 1163 542 1297">番組AA</div> <div data-bbox="574 1163 617 1297">番組AB</div> <div data-bbox="649 1402 691 1556">10/24(金)</div> <div data-bbox="649 1163 691 1297">番組BA</div>	<div data-bbox="493 596 535 877">視聴されていない</div> <div data-bbox="574 596 617 877">視聴されていない</div> <div data-bbox="649 596 691 877">視聴されていない</div>

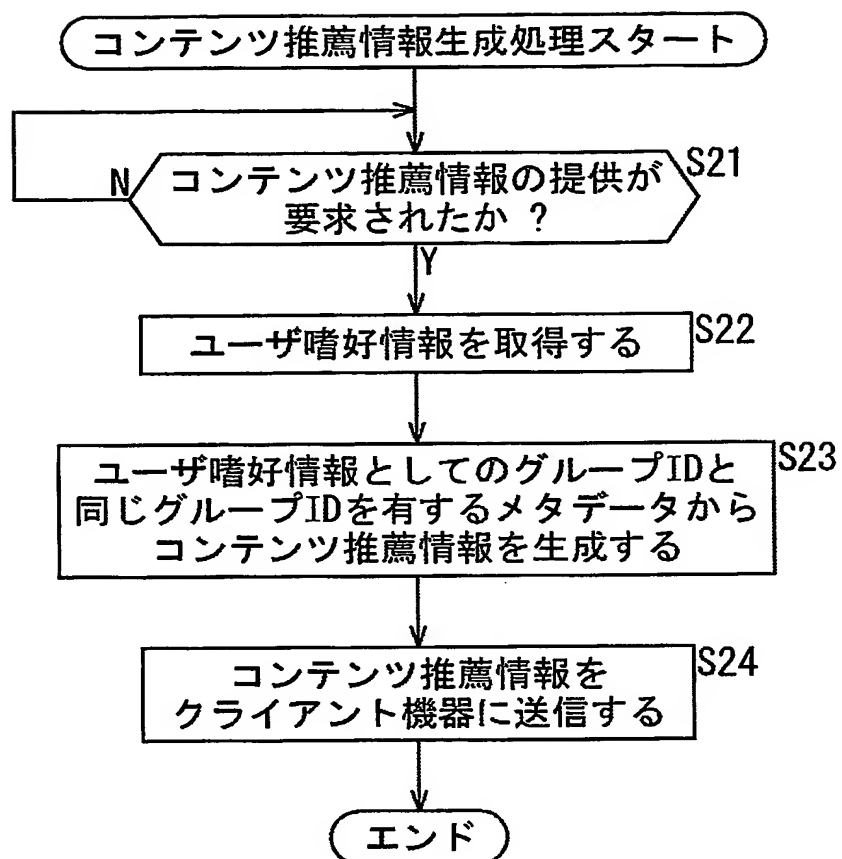
【図 14】

図14

配信リスト 「8ch, 20:00~21:00」	利用履歴 「8ch, 20:00~21:00」
<div data-bbox="505 1402 545 1556">10/25(土)</div> <div data-bbox="505 1192 545 1297">番組 A</div>	<div data-bbox="500 596 540 877">視聴されていない</div>
<div data-bbox="578 1402 618 1556">10/24(金)</div> <div data-bbox="578 1192 618 1297">番組 B</div>	<div data-bbox="573 701 613 877">視聴された</div>
<div data-bbox="654 1192 695 1297">番組 C</div>	<div data-bbox="649 701 690 877">視聴された</div>
<div data-bbox="732 1402 773 1556">10/23(木)</div> <div data-bbox="732 1192 773 1297">番組 D</div>	<div data-bbox="727 701 768 877">視聴された</div>

【図 15】

図15



【図 16】

図16

57 {

あなたの好きな「時間帯」でお勧めします!

8ch, 20:00~21:00 グループID= ××××

・番組AAAのタイトル

10ch, 19:00~20:00 グループID= ××××

・番組BBAのタイトル

【図17】

図17

57 {

あなたの好きな「ジャンル」,「出演者」でお勧めします

バラエティ, タレントC グループID=xxxxx

ニュース, タレントD グループID=xxxxx

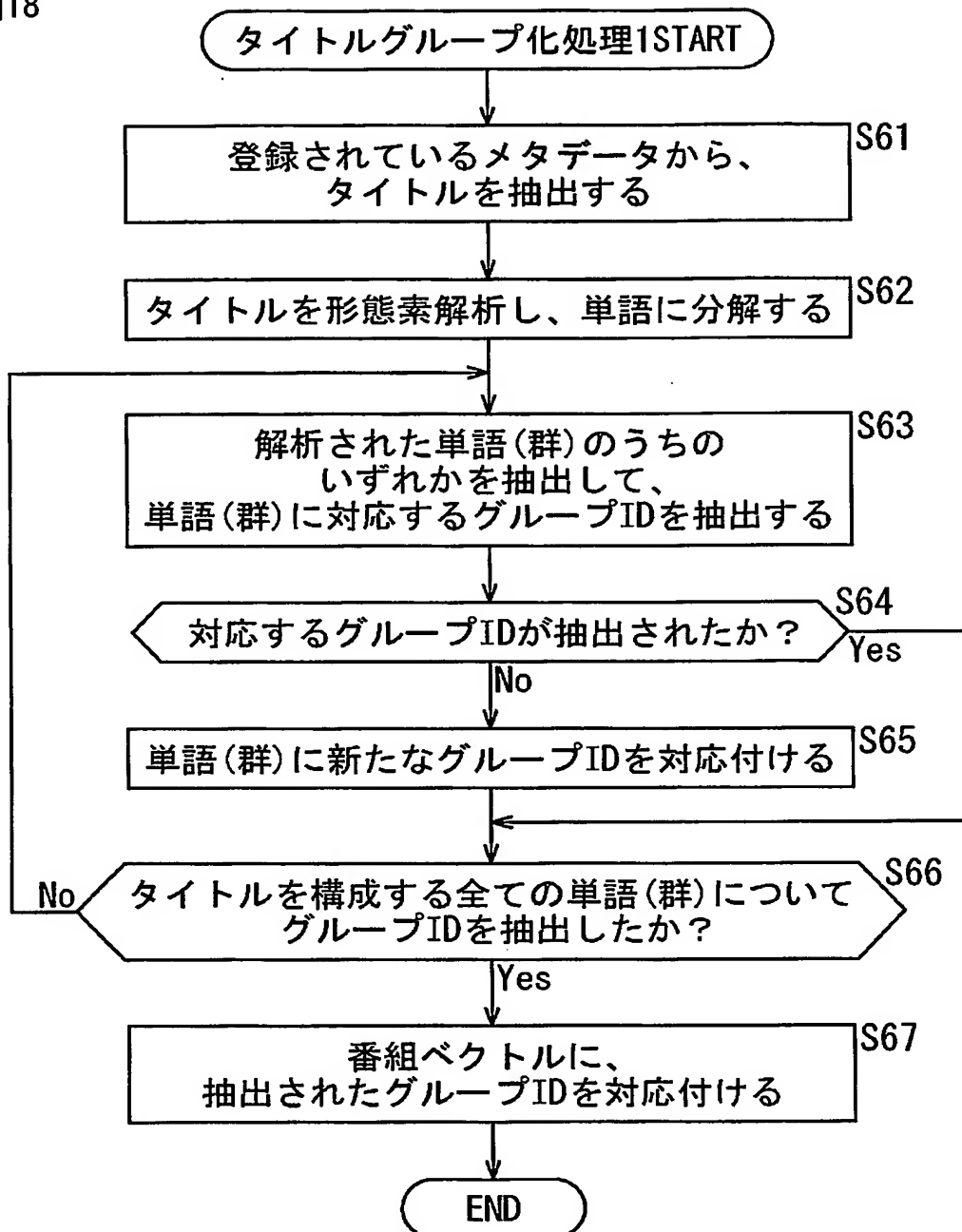
・番組CCAのタイトル

・番組CCBのタイトル

⋮

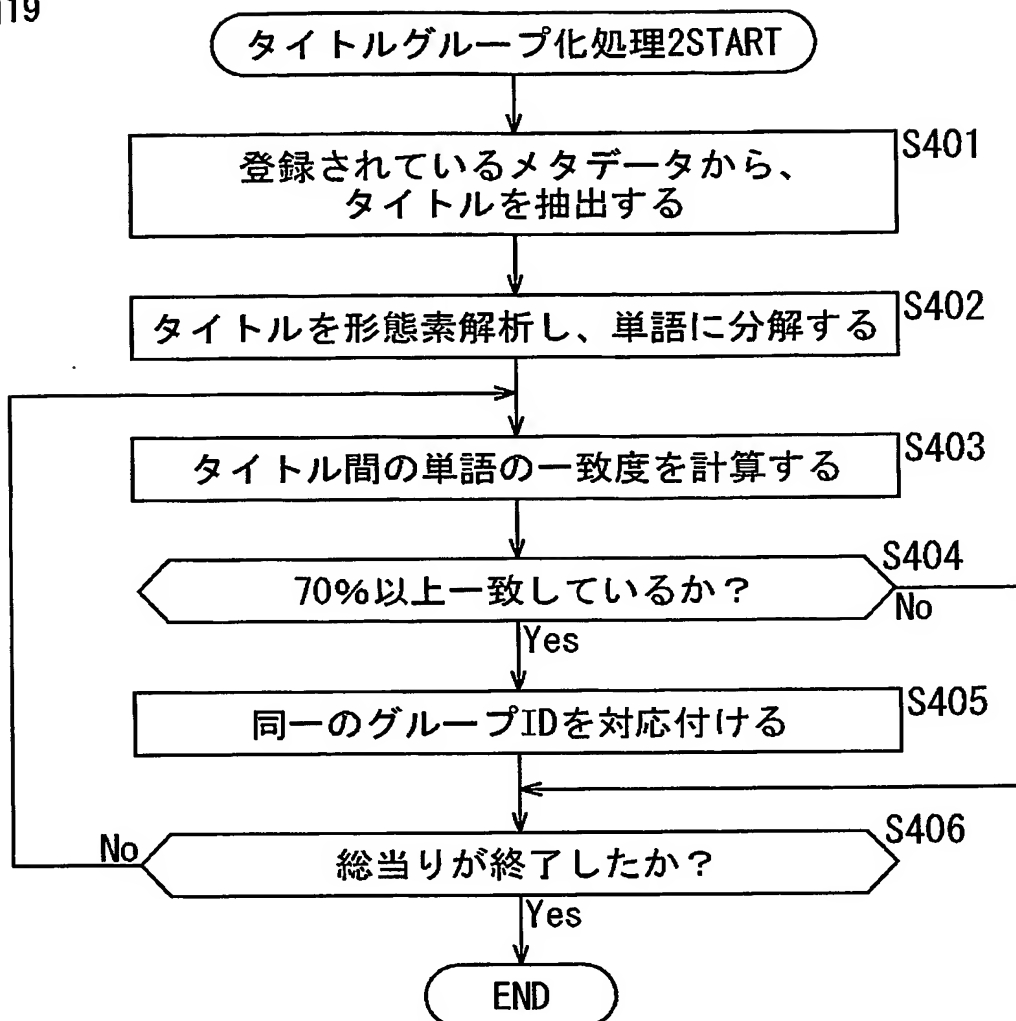
【図 18】

図18



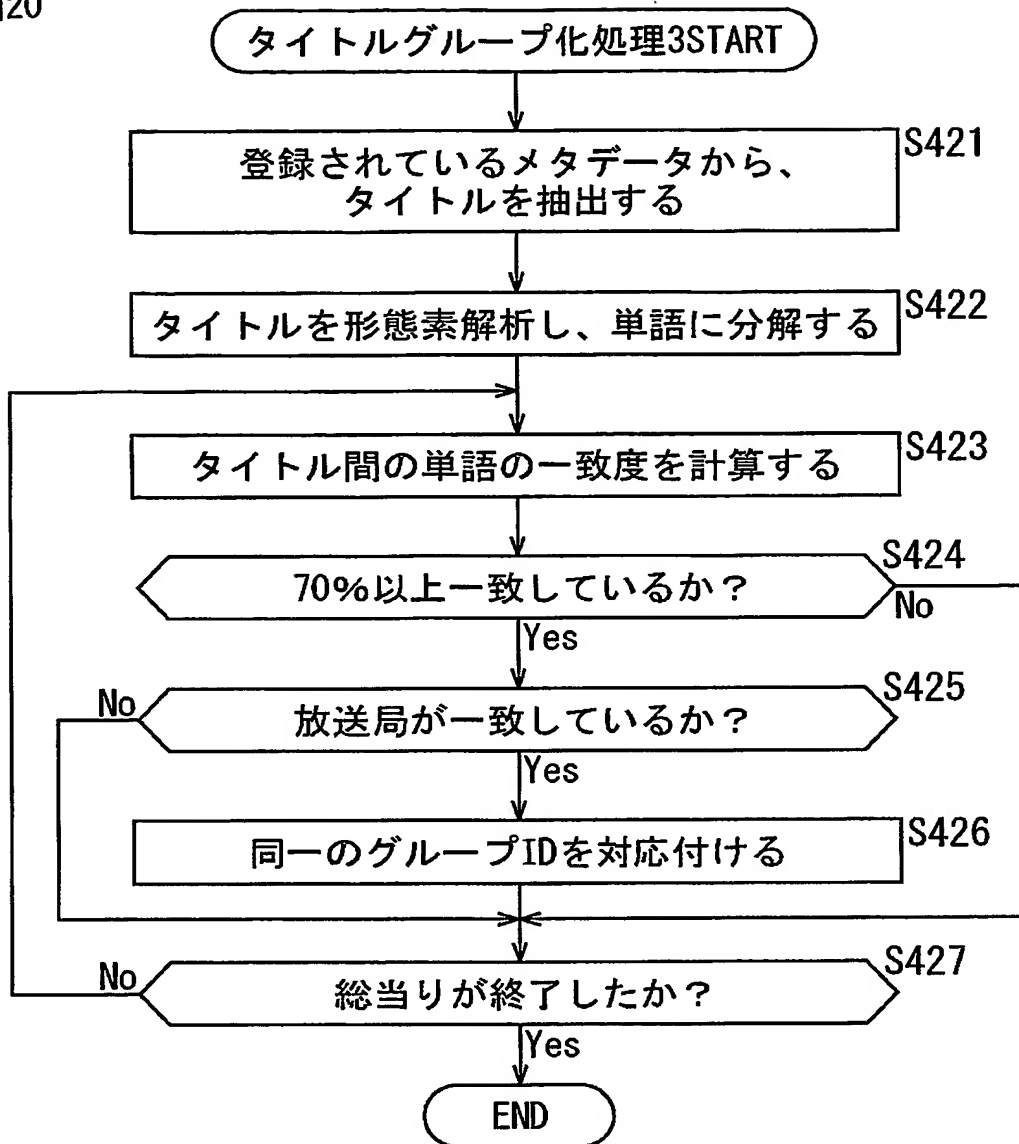
【図 19】

図19



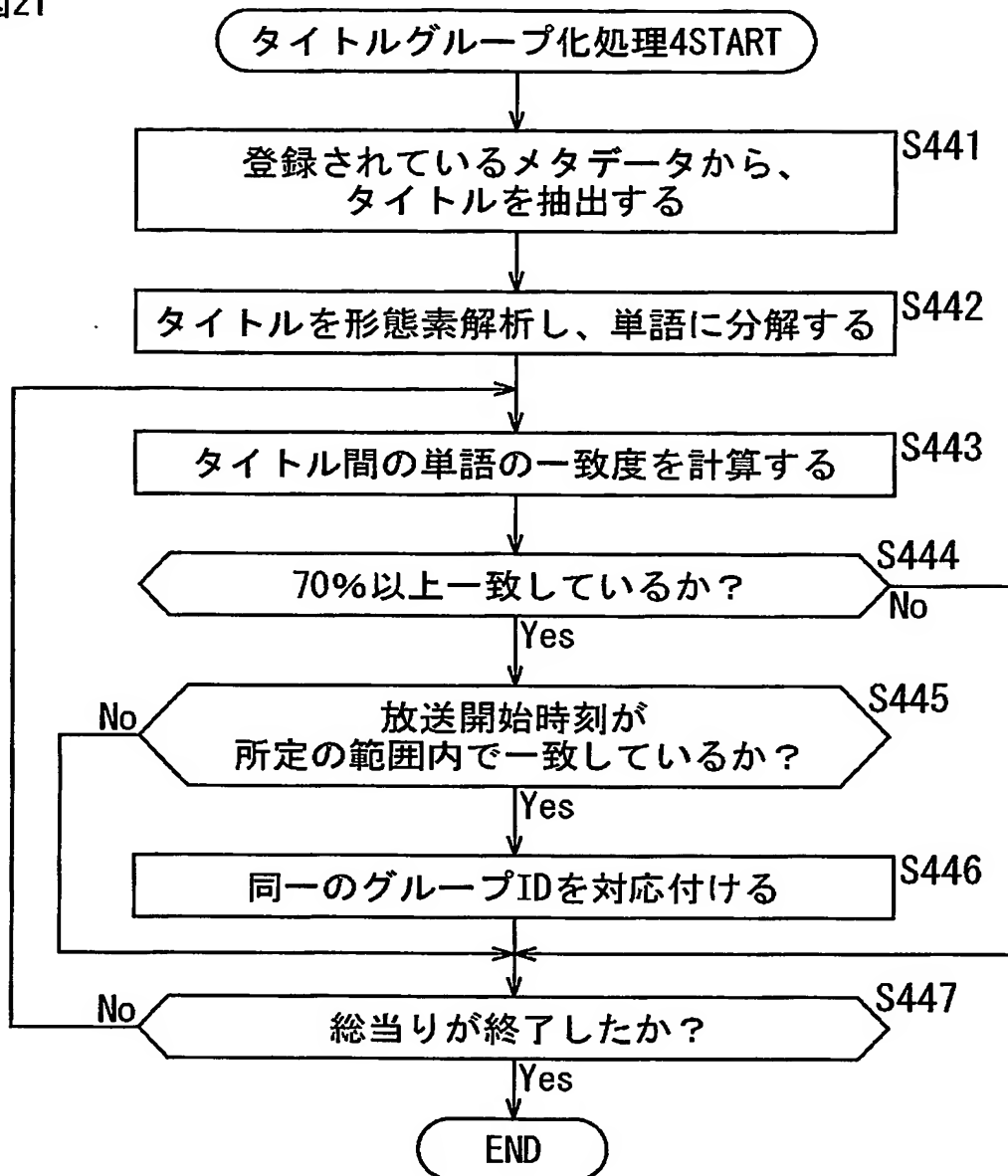
【図 20】

図20



【図 21】

図21



【書類名】 要約書

【要約】

【課題】 グループ単位の利用頻度を利用してユーザ嗜好情報を生成する。

【解決手段】 ステップS2において、メタデータの所定の項目からなるグループ化項目におけるグループ毎に利用頻度が算出される。ステップS3において、閾値以上の利用頻度のグループが検出される。ステップS4において、検出されたグループに属するコンテンツの利用状態が所定の要件を満たすか否かが確認され、ステップS5においてその条件を満たすグループが検出される。ステップS6において、検出されたグループに付されたグループIDがユーザ嗜好情報とされる。ユーザ嗜好情報としてのグループIDと同じグループIDを有するコンテンツがユーザに推薦される。

【選択図】 図9

特願 2002-361540

出願人履歴情報

識別番号

[000002185]

1. 変更年月日

1990年 8月30日

[変更理由]

新規登録

住 所

東京都品川区北品川6丁目7番35号

氏 名

ソニー株式会社